Paper alert

Chosen by Lynne Regan[1], Luis Serrano[2], Andrej Šali[3], Amnon Horovitz[4] and Charles Wilson[5]

A selection of interesting papers published in recent months in those major journals most likely to report significant results in protein and RNA folding and design.

Addresses: [1]Department of Molecular Biophysics and Biochemistry, Yale University, 266 Whitney Avenue, Box 6666, New Haven, CT 06520, USA. [2]EMBL, Meyerhofstrasse 1, Heidelberg D-69012, Germany. [3]Box 270, Rockefeller University, 1230 York Avenue, New York, NY 10021-6399, USA. [4]Department of Structural Biology, Weizmann Institute, IL-76100, Rehovot, Israel. [5]Department of Biology and Center for the Molecular Biology of RNA, University of California, Santa Cruz, CA 95064, USA.

Folding & Design 1997, **2**:R81–R84

☐ **Identification and structural characterization of the ATP/ADP-binding site in the Hsp90 molecular chaperone.** Chrisostomos Prodromou, S Mark Roe, Ronan O'Brien, John E Ladbury, Peter W Piper and Laurence H Pearl. *Cell* **90,** 65–75.

Hsp90 molecular chaperones in eukaryotic cells play essential roles in the folding and activation of a range of client proteins involved in cell cycle regulation, steroid hormone responsiveness, and signal transduction. The biochemical mechanism of Hsp90 is poorly understood, and the involvement of ATP, in particular, is controversial. Crystal structures of complexes between the N-terminal domain of the yeast Hsp90 chaperone and ADP/ATP unambiguously identify a specific adenine nucleotide binding site homologous to the ATP-binding site of DNA gyrase B. This site is the same as that identified for the antitumor agent geldanamycin, suggesting that geldanamycin acts by blocking the binding of nucleotides to Hsp90 and not the binding of incompletely folded client polypeptides as previously suggested. These results finally resolve the issue of the direct involvement of ATP in Hsp90 function.
11 July 1997. *Cell.*

☐ **Coupling backbone flexibility and amino acid sequence selection in protein design.** Alyce Su and Stephen L Mayo. *Protein Sci.* **4,** 630–634.

Using a protein design algorithm that considers sidechain packing quantitatively, the effect of explicit backbone motion on the selection of amino acids in protein design was assessed in the core of the streptococcal protein Gβ1 domain. Concerted backbone motion was introduced by varying the super-secondary structure parameter values of Gβ1. The stability and structural flexibility of seven of the redesigned proteins were determined experimentally and showed that core variants containing as many as six of 10 possible mutations retain native-like properties. This result demonstrates that backbone flexibility can be combined explicitly with amino acid sidechain selection and that the selection algorithm is sufficiently robust to tolerate perturbations as large as 15% of Gβ1's native supersecondary structure parameter values.
August 1997. *Protein Science.*

☐ **A residue-specific NMR view of the non-cooperative unfolding of a molten globule.** Brenda A Schulman, Peter S Kim, Christopher M Dobson and Christina Redfield. *Nat. Struct. Biol.* **4,** 630–634.

Molten globules are partially folded forms of proteins that are thought to be general intermediates in protein folding. Nonetheless, there is limited structural information about such species because they possess conformational heterogeneity and complex dynamical properties that lead to extreme line broadening in NMR spectra. The authors use a 2D NMR approach that overcomes this difficulty by detecting the unfolding of individual residues in a molten globule in increasing concentrations of denaturant. The results show that the structure in the low pH form of α-lactalbumin is not formed cooperatively. Moreover, a core region remains collapsed under extremely denaturing conditions, even when the majority of the polypeptide chain is completely unfolded. These results support a model for protein folding in which the core provides a template for correct assembly of the remainder of the structure.
August 1997. *Nature Structural Biology.*

☐ **Dissecting contributions to the thermostability of *Pyrococcus furiosus* rubredoxin: β-sheet chimeras.** Marly K Eidsness, Kimberly A Richie, Amy E Burden, Donald M Kurtz Jr and Robert A Scott. *Biochemistry* **36,** 10406–10413.

The contributions to thermostability of interactions within the β-sheet region of rubredoxins were investigated by examining proteins in which β-strand sequences of rubredoxins from the hyperthermophilic archaeon *Pyrococcus furiosus* (Pf) and the mesophilic bacterium *Clostridium pasteurianum* (Cp) were interchanged. The thermostabilities of the chimeric rubredoxins were assessed by monitoring the decay of the visible absorbance at 490 nm and of the far-UV CD versus time at 92°C. The chimeric rubredoxins Pf15|Cp47|Pf (Pf rubredoxin residues 2–15 and 48–54 and Cp rubredoxin residues 16–47) and Cp15|Pf47|Cp were both found to be far less thermostable than wild-type Pf rubredoxin, indicating that neither the β-sheet residues (2–7, 10–15, and 48–53) nor the 'core residues' (16–47) of Pf rubredoxin independently confer Pf rubredoxin-like thermostability. However, the chimeric Pf47|Cp exhibits thermostability close to that of wild-type Pf rubredoxin, suggesting that Pf rubredoxin-like thermostability is conferred by interactions of β-sheet strands 1 and 2 (residues 2–15) together with Pf core residues.
August 1997. *Biochemistry.*

☐ **Absence of a stable intermediate on the folding pathway of protein A.** Yawen Bai, Afshin Karimi, H Jane Dyson and Peter E Wright. *Protein Sci.* **6,** 1449–1457.

The B-domain of protein A has one of the simplest protein topologies, a three-helix bundle, and its folding has been studied as a model for elementary steps in the folding of larger proteins. Earlier studies suggested that folding might occur by way of a helical hairpin intermediate. Equilibrium hydrogen exchange measurements indicate that the C-terminal helical hairpin could be a potential folding intermediate. Kinetic refolding experiments were performed using stopped-flow CD and NMR hydrogen/deuterium exchange pulse labeling. Folding of the entire molecule is essentially complete within the 6 ms dead time of the quench-flow apparatus, indicating that the intermediate, if formed, progresses rapidly to the final folded state. Site-directed mutagenesis of the isoleucine residue at position 16 was used to generate a variant protein containing tryptophan (the 116W mutant). The formation of the putative folding intermediate was expected to be favored in this mutant at the expense of the native folded form, due to predicted unfavorable steric interactions of the bulky tryptophan sidechain in the folded state. The 116W mutant refolds completely within the dead time of a stopped-flow fluorescence experiment. These results explicitly demonstrate that the B-domain of protein A folds extremely rapidly by an apparent two-state mechanism without formation of stable partly folded intermediates.
July 1997. *Protein Science.*

☐ **Designed protein pores as components for biosensors.** Orit Braha, Barbara Walker, Stephen Cheley, John J Kasianowicz, Langzhou Song, J Eric Gouaux and Hagan Bayley. *Chem. Biol.* **4,** 497–505.

There is a pressing need for new sensors that can detect a variety of analytes, ranging from simple ions to complex compounds and even microorganisms. The devices should offer sensitivity, speed, reversibility and selectivity. Given these criteria, protein pores, remodeled so that their transmembrane conductances are modulated by the association of specific analytes, are excellent prospects as components of biosensors. Structure-based design and a separation method that employs targeted chemical modification have been used to obtain a heteromeric form of the bacterial pore-forming protein staphylococcal α-hemolysin, in which one of the seven subunits contains a binding site for a divalent metal ion, M(II), which serves as a prototypic analyte. The single-channel current of the heteromer in planar bilayers is modulated by nanomolar Zn(II). Other M(II)s modulate the current and produce characteristic signatures. Engineered pores have several advantages as potential sensor elements: sensitivity is in the nanomolar range; analyte binding is rapid (diffusion limited in some cases) and reversible; strictly selective binding is not required because single-channel recordings are rich in information; and for a particular analyte, the dissociation rate constant, the extent of channel block and the voltage-dependence of these parameters are distinguishing. The approach described here can be generalized for additional analytes.
July 1997. *Chemistry & Biology.*

☐ **Tertiary structure of RBD2 and backbone dynamics of RBD1 and RBD2 of the human U1A protein determined by NMR spectroscopy.** Jirong Lu and Kathleen B Hall. *Biochemistry* **36,** 10393–10405.

The human U1A protein has two putative RNA-binding domains, one in the N-terminal region of the protein (RBD1) and the other at the C-terminal end (RBD2). RBD1 binds tightly and specifically to one of the stem loops of the U1 snRNA, as well as to its own 3′-UTR. In contrast, RBD2 does not appear to associate with any RNA. The two domains share 25% amino acid identity, and both have the same βαβ–βαβ secondary structure fold. In this work, $^{13}C/^{15}N/^{1}H$ multi-dimensional NMR methods were used to obtain sidechain assignments for RBD2, and then the tertiary structure was calculated using a distance geometry/simulated annealing algorithm that employs pairwise Gaussian metrization. RBD2 is shown to fold into an α/β sandwich with a four-stranded antiparallel β-sheet, which is the typical global topology of these domains. The ^{15}N backbone dynamics of these two structurally homologous RBDs are significantly different. Conformational exchange observed in RBD1, which is absent in RBD2, may correlate with the mechanism of RNA binding.
26 August 1997. *Biochemistry.*

☐ **ESCHER: a new docking procedure applied to the reconstruction of protein tertiary structure.** G Ausiello, G Cesareni and M Helmer-Citterich. *Proteins* **28,** 556–567.

Evaluation of surface complementarity, hydrogen bonding, and electrostatic interaction in molecular recognition (ESCHER) is a new docking procedure consisting of three modules that work in series. The first module evaluates the geometric complementarity and produces a set of rough solutions for the docking problem. The second module identifies molecular collisions within those solutions. The third evaluates their electrostatic complementarity. We describe the algorithm and its application to the docking of cocrystallized protein domains and unbound components of protein–protein complexes. Furthermore, ESCHER has been applied to the reassociation of secondary and supersecondary structure elements. The possibility of applying a docking method to the problem of protein structure prediction is discussed.
August 1997. *Proteins: Structure, Function, and Genetics.*

☐ **Simultaneous and coupled energy optimization of homologous proteins: a new tool for structure prediction.** Chen Keasar, Ron Elber and Jeffrey Skolnick. *Fold. Des.* **2,** 247–259.

Homology-based modeling and global optimization of energy are two complementary approaches to prediction of protein structures. A combination of the two approaches is proposed in which a novel component is added to the energy and forces similarity between homologous proteins. The combination was tested for two families: pancreatic hormones and homeodomains. The simulated lowest-energy structure of the pancreatic hormones is a reasonable approximation to the native fold. The lowest-energy structure of the homeodomains has

80% of the native contacts, but the helices are not packed correctly. The fourth lowest energy structure of the homeodomains has the correct helix packing (RMS 5.4 Å and 82% of the correct contacts). Optimizations of a single protein of the family yield considerably worse structures. Use of coupled homologous proteins in the search for the native fold is more successful than the folding of a single protein in the family.
17 July 1997. *Folding & Design.*

☐ **Pfam: a comprehensive database of protein domain families based on seed alignments.** Erik LL Sonnhammer, Sean R Eddy and Richard Durbin. *Proteins* **28,** 405–420.

Databases of multiple sequence alignments are a valuable aid to protein sequence classification and analysis. One of the main challenges when constructing such a database is to simultaneously satisfy the conflicting demands of completeness on the one hand and quality of alignment and domain definitions on the other. The latter properties are best dealt with by manual approaches, whereas completeness in practice is amenable only to automatic methods. The authors present a database based on hidden Markov model profiles which combines high quality and completeness. By using this database, *Pfam*, a large number of previously unannotated proteins from the *Caenorhabditis elegans* genome project were classified. The authors have also identified many novel family memberships in known proteins, including new kazal, fibronectin type III, and response regulator receiver domains.
July 1997. *Proteins: Structure, Function, and Genetics.*

☐ **Correlated mutations contain information about protein–protein interaction.** Florencio Pazos, Manuela Helmer-Citterich, Gabriele Ausiello and Alfonso Valencia. *J. Mol. Biol.* **271,** 511–523.

Many proteins have evolved to form specific molecular complexes and the specificity of this interaction is essential for their function. The network of the necessary interresidue contacts must consequently constrain the protein sequences to some extent. In other words, the sequence of an interacting protein must reflect the consequence of this process of adaptation. It is reasonable to assume that the sequence changes accumulated during the evolution of one of the interacting proteins must be compensated by changes in the other. The authors apply a method for detecting correlated changes in multiple sequence alignments to a set of interacting protein domains and show that positions where changes occur in a correlated fashion in the two interacting molecules tend to be close to the protein–protein interfaces. This leads to the possibility of developing a method for predicting contacting pairs of residues from the sequence alone. Such a method would not need the knowledge of the structure of the interacting proteins, and hence would be both radically different and more widely applicable than traditional docking methods. The authors indeed demonstrate that the information about correlated sequence changes is sufficient to single out the right interdomain docking solution amongst many wrong alternatives of two-domain proteins. The same approach is also used here in one case (haemoglobin) where the authors attempt to predict the interface of two different proteins rather than two protein domains. Finally, they report a prediction about the interdomain contact regions of the heat-shock protein Hsc70 based only on sequence information.
29 August 1997. *Journal of Molecular Biology.*

☐ **Effectiveness of correlation analysis in identifying protein residues undergoing correlated evolution.** DD Pollock and WR Taylor. *Protein Eng.* **10,** 647–657.

Various methods for detecting correlation between sites were evaluated by ascertaining their ability to discriminate positively correlated sites from background correlation at randomly evolved sites. A model for generating pairwise correlations of different degrees is also described. An assortment of physicochemical vectors and similarity and difference matrices were used to discriminate correlated change. There was little difference in effectiveness between the different matrices, but there were significant differences between the matrices and the physicochemical vectors. It is shown that all methods investigated exhibit significant inability to screen out background correlation, particularly in the presence of phylogenetic relatedness between the sequences. Methods using the matrices are unable to distinguish positively correlated from negatively correlated, or compensatory, replacements.
June 1997. *Protein Engineering.*

☐ **The chaperonin cycle cannot substitute for prolyl isomerase activity, but GroEL alone promotes productive folding of a cyclophilin-sensitive substrate to a cyclophilin-resistant form**. O von Ahsen, M Tropschug, N Pfanner and J Rassow. *EMBO J.* **15,** 4568–4578.

The chaperonin GroEL and the peptidyl–prolyl *cis–trans* isomerase cyclophilin are major representatives of two distinct cellular systems that help proteins to adopt their native three-dimensional structure: molecular chaperones and folding catalysts. Little is known about whether and how these proteins cooperate in protein folding. In this study, the authors have examined the action of GroEL and cyclophilin on a substrate protein in two distinct prolyl isomerization states. Their results indicate that GroEL binds the same substrate in different prolyl isomerization states; GroEL–ES does not promote prolyl isomerizations, but even retards isomerizations; cyclophilin cannot promote the correct isomerization of prolyl bonds of a GroEL-bound substrate, but acts sequentially after release of the substrate from GroEL; a denatured substrate with all-native prolyl bonds is delayed in folding by cyclophilin due to isomerization to non-native prolyl bonds (a substrate that has proceeded in folding beyond a stage where it can be bound by GroEL is still sensitive to cyclophilin); if a denatured cyclophilin-sensitive substrate is first bound to GroEL, however, productive folding to a cyclophilin-resistant form can be promoted, even without GroES. The authors conclude that

GroEL and cyclophilin act sequentially and exert complementary functions in protein folding.
01 August 1997. *The EMBO Journal.*

☐ **The crystal structure of the asymmetric GroEL–GroES–$(ADP)_7$ chaperonin complex.** Z Xu, AL Horwich and PB Sigler. *Nature* **388,** 741–791.
Chaperonins assist protein folding with the consumption of ATP. They exist as multi-subunit protein assemblies comprising rings of subunits stacked back to back. In *Escherichia coli*, asymmetric intermediates of GroEL are formed with the co-chaperonin GroES and nucleotides bound only to one of the seven-subunit rings (the *cis* ring) and not to the opposing ring (the *trans* ring). The structure of the GroEL–GroES–$(ADP)_7$ complex reveals how large *en bloc* movements of the *cis* rings intermediate and apical domains enable bound GroES to stabilize a folding chamber with ADP confined to the *cis* ring. Elevation and twist of the apical domains double the volume of the central cavity and bury hydrophobic peptide-binding residues in the interface with GroES, as well as between GroEL subunits, leaving a hydrophilic cavity lining that is conducive to protein folding. An inward tilt of the *cis* equatorial domain causes an outward tilt in the *trans* ring that opposes the binding of a second GroES. When combined with new functional results, this negative allosteric mechanism suggests a model for an ATP-driven folding cycle that requires a double toroid.
21 August 1997. *Nature.*

☐ **Recombination of protein domains facilitated by co-translational folding in eukaryotes.** William J Netzer and F Ulrich Hartl. *Nature* **388,** 343–349.
The evolution of complex genomes requires that new combinations of pre-existing protein domains successfully fold into modular polypeptides. During eukaryotic translation, model two-domain polypeptides fold efficiently by sequential and co-translational folding of their domains. In contrast, folding of the same proteins in *E. coli* is posttranslational and leads to intramolecular misfolding of concurrently folding domains. Sequential domain folding in eukaryotes may have been critical in the evolution of modular polypeptides, by increasing the probability that random gene-fusion events resulted in immediately foldable protein structures.
24 July 1997. *Nature.*

☐ ***In vivo* observation of polypeptide flux through the bacterial chaperonin system.** Karla L Ewalt, Joseph P Hendrick, Walid A Houry and F Ulrich Hartl. *Cell* **90,** 491–500.
The quantitative contribution of chaperonin GroEL to protein folding in *E. coli* was analyzed. A diverse set of newly synthesized polypeptides, predominantly between 10 and 55 kDa, interacts with GroEL, accounting for 10–15% of all cytoplasmic protein under normal growth conditions, and for 30% or more upon exposure to heat stress. Most proteins leave GroEL rapidly within 10–30 s. The authors distinguish three classes of substrate proteins: proteins with a chaperonin-independent folding pathway; proteins (more than 50% of total) with an intermediate chaperonin dependence for which normally only a small fraction transits GroEL; and a set of highly chaperonin-dependent proteins, many of which dissociate slowly from GroEL and probably require sequestration of aggregation-sensitive intermediates within the GroEL cavity for successful folding.
08 August 1997. *Cell.*

☐ **Structural basis of allosteric changes in the GroEL mutant Arg197*Ala.** Helen E White, Shaoxia Chen, Alan M Roseman, Ofer Yifrach, Amnon Horovitz and Helen R Saibil. *Nat. Struct. Biol.* **4,** 690–693.
Combined kinetic and cryo-electron microscopy analysis of the Arg197Ala mutant of GroEL provides insight into the allosteric switching of GroEL, which is at the heart of the chaperonin mechanism.
September 1997. *Nature Structural Biology.*

Current articles of interest to readers of *Folding & Design*

A selection of dispatches and other review articles published in other journals from Current Biology Ltd. Selected by the staff of *Folding & Design*.

Touring the landscapes: partially folded proteins examined by hydrogen exchange [Minireview]. Aaron K Chamberlain and Susan Marqusee (1997). *Structure* **5,** 859–863.

Examining the conformational dynamics of macromolecules with time-resolved synchrotron X-ray 'footprinting' [Ways & Means]. Mark R Chance, Bianca Sclavi, Sarah A Woodson and Michael Brenowitz (1997). *Structure* **5,** 865–869.

Molecular dynamics simulations [Primer]. Bruce Tidor (1997). *Current Biology* **7,** R480–R482.

Molecular chaperones: Avoiding the crowds [Dispatch]. R John Ellis (1997). *Current Biology* **7,** R531–R533.

Seven reviews on aspects of **Engineering and design** in the August issue (No 4) of *Current Opinion in Structural Biology* (1997), **7,** 455–508. Edited by William F DeGrado and Björn O Nilsson.

Sixteen reviews on aspects of **Protein engineering** in the August issue (No 4) of *Current Opinion in Biotechnology* (1997), **7,** 397–508. Edited by Birgit A Helm and Eduardo A Padlan.